**Geostatistics Software User's
Manual for the Geosciences
Research and Engineering
Department PDP 11/70
Computer**

J. L. Devary
P. J. Mitchell

W. A. Rice
D. W. Damschen

December 1983

Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Operated for the U.S. Department of Energy
by Battelle Memorial Institute

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America
Available from
National Technical Information Service
United States Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161

NTIS Price Codes
Microfiche A01

Printed Copy

Pages	Price Codes
001-025	A02
026-050	A03
051-075	A04
076-100	A05
101-125	A06
126-150	A07
151-175	A08
176-200	A09
201-225	A010
226-250	A011
251-275	A012
276-300	A013

GEOSTATISTICS SOFTWARE USER'S MANUAL
FOR THE GEOSCIENCES RESEARCH AND
ENGINEERING DEPARTMENT PDP 11/70
COMPUTER

J. L. Devary
P. J. Mitchell
W. A. Rice
D. W. Damschen

December 1983

Prepared for
the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Richland, Washington 99352

SUMMARY

Geostatistics and kriging are statistical techniques that can be used to estimate a hydrologic parameter surface from spatially distributed data. The techniques produce statistically optimum estimates of a surface and can aid in pinpointing areas of greatest need for data. This document describes the use of geostatistical software routines that have been implemented at Pacific Northwest Laboratory's (PNL's) PDP 11/70 computer of the Geosciences Research and Engineering Department for use with the Hanford ground water monitoring data base of the Ground-Water Surveillance Program sponsored by U.S. Department of Energy. The geostatistical routines, originally developed by J. P. Hughes for his masters thesis at the University of Washington and subsequently modified by W. A. Rice (PNL) on the 11/70, were modified to improve input/output file efficiency and to accept larger data sets. Two data processing routines were also developed for retrieving and sorting data into a form compatible with the geostatistical routines.

Chapter 2.0 contains brief descriptions of the geostatistical theory involved in FINDK, KHREG, and KRIEST. Chapter 3.0 contains complete instructions to execute any of the four routines as well as sample input and output for each routine.

CONTENTS

SUMMARY.....	iii
1.0 INTRODUCTION.....	1.1
2.0 GEOSTATISTICAL THEORY.....	2.1
2.1 DRIFT.....	2.3
2.2 GENERALIZED INCREMENTS AND GENERALIZED COVARIANCES.....	2.6
2.3 KRIGING.....	2.8
2.4 DRIFT ORDER IDENTIFICATION.....	2.12
2.5 GENERALIZED COVARIANCE IDENTIFICATION.....	2.14
2.6 REFERENCES.....	2.18
3.0 USER INSTRUCTIONS AND SAMPLE CASES.....	3.1
3.1 KRGRTV.....	3.1
3.2 KRIGTRAN.....	3.5
3.3 FINDK.....	3.7
3.4 KHREG.....	3.11
3.5 KRIEST.....	3.17
3.6 PLTWELL.....	3.20
4.0 BIBLIOGRAPHY.....	4.1
DISTRIBUTION.....	Distr-1

FIGURES

2.1	Common Variogram Models for Scaling Factor $w = 1.0$ (from Delhomme 1976).....	2.4
2.2	Linear Variogram with Nugget Effect, C.....	2.5
3.1	Sample Listing of KRGRTV Output (T0667.DAT).....	3.3
3.2	Sample Listing of KRIGTRAN Output (T0667.TMP).....	3.6
3.3	Sample Output of FINDK.....	3.8
3.4	Sample Run of KHREG.....	3.13
3.5	Sample Run of KRIST.....	3.21
3.6	Sample Run of PLTWELL.....	3.28
3.7	Plot from PLTWELL with Tritium Contours, for 1967.....	3.29

TABLES

3.1	Contaminant Code for the CIRMIS Data Base Management System.....	3.4
-----	--	-----

1.0 INTRODUCTION

Geostatistics and kriging are statistical techniques that can be used to estimate a hydrologic parameter surface from spatially distributed data. The techniques produce statistically optimum estimates of a surface and can aid in pinpointing areas of greatest need for data. This document describes the use of geostatistical software routines that have been implemented on the Geosciences Research and Engineering Department PDP 11/70 computer for use with the ground water monitoring data base of the Ground-Water Surveillance Program at PNL. The geostatistical routines were originally developed by J. P. Hughes for his masters thesis at the University of Washington and subsequently modified by W. A. Rice (PNL) on the 11/70 computer for masters thesis work. The final versions of these routines, described herein, were modified to improve input/output file efficiency and to accept larger data sets. Two data processing routines were also developed for retrieving and sorting data into a form compatible with the geostatistical routines.

Kriging is a statistical interpolation technique that permits the estimation of a hydrologic parameter surface from discrete, spatially distributed data. (Potentiometric, hydraulic conductivity, transmissivity, topographic, and contaminant plume data have been kriged with great success.) Thus, kriging can interpolate irregularly spaced data values to a regular rectangular grid that may be used as the data base for a contour plotting program (e.g., CNT) or for ground water modeling. Areas in greatest need of data (e.g., additional sampling points) can be pinpointed by locating areas of greatest estimation errors or uncertainty.

Statistically, kriging is an unbiased, minimum variance estimator. This means that not only the average estimation (interpolation) error is zero, but also the magnitude of error is expected to be less than errors associated with other estimation procedures. For example, standard least squares regression techniques ignore any spatial correlation or continuity present in the observed data, thereby increasing the magnitude of errors. The use of the most efficient statistical algorithms to estimate surfaces is necessary to extract the

maximum amount of information from monitoring well data, which is often very expensive to collect.

Besides providing a precise estimate of the parameter surface, kriging also permits the statistical quantification of error associated with the estimated surface. This provides the means to investigate the validity of the observed data, evaluate measurement error, and perform sensitivity studies.

In order to krig a hydrologic parameter surface, it is necessary to know the drift and covariance structure of the observed data. The drift is the slowly varying, continuous component of the observed data base. For example, the drift associated with potentiometric data is generally a linear function of spatial position with coefficients equal to the average or mean hydraulic gradient for the region. The covariance, or more precisely the generalized covariance, describes stochastically the highly variable ("noisy" or irregular), seemingly random component of the observed data that is superimposed upon the drift. The proper choice of covariance can be used to model highly variable hydraulic conductivity data as well as smooth potentiometric or contaminant plume data.

The following geostatistics routines have been implemented on the Geosciences Research and Engineering PDP 11/70 computer system:

1. FINDK A program that selects the order of drift for the observed data base. The drift is modeled as a polynomial function of spatial position; the order of drift indicates the growth of the slowly varying, continuous component of the data. Order = 0 indicates no trend in data; order = 1 indicates a linear trend; order = 2 indicates a quadratic trend.
Input: Terminal, Observed data base file
Output: Lineprinter
2. KHREG A program that selects the generalized covariance function for the observed data base given the order of drift. Typically,

FINDK is run before KHREG. The generalized covariance is modeled as a polynomial.

Input: Terminal, Observed data base file

Output: Lineprinter

3. KRIEST A program that krige estimates of a parameter surface on a regular rectangular grid from the observed data base given the order of drift and the generalized covariance function. Typically, FINDK and KHREG are run before KRIEST.
Input: Terminal, Observed data base file
Output: Line printer, Krige surface file, Standard deviation (of kriging error) file
4. PLTWELL A program to generate a plot of the observed data and contour plots from KRIEST. PLTWELL permits the superposition of krige surface contours on the observed data base.
Input: Terminal, Observed data base file
Output: Lineprinter, Observed data base plot file, CNT control file for contour plot of krige surface.

Two additional data processing routines were developed to sort and transform Hanford ground water data into a form compatible with FINDK, KHREG, or KRIEST.

5. KRGRTV A program to retrieve contaminant, hydrograph, or temperature data from the data base for a particular time period and a particular rectangular subregion of the Hanford Reservation.
Input: Terminal, CIRMIS data base file
Output: Lineprinter, Well data file 1
6. KRIGTRAN A program that scales the data output from KRGRTV so that FINDK, KHREG, KRIEST, and PLTWELL will not encounter numerical overflow problems. KRIGTRAN utilizes only the maximal data value if replicate measurements exist at any one well location.
Input: Terminal, Well data file 1
Output: Lineprinter, Well data file 2

This concludes the introduction of the geostatistical software package. Chapter 2.0 contains brief descriptions of the geostatistical theory involved in FINDK, KHREG, and KRIEST. The user not interested in the mathematical theory can skip directly to Chapter 3.0 and find complete instructions to execute any of the four routines; sample input and output for each routine is included.

2.0 GEOSTATISTICAL THEORY

This chapter details the important points of the theory of regionalized variables, develops the theoretical basis for kriging, and reviews the methods that have been used for drift and generalized covariance estimation. The development follows Olea (1975) and includes results from Delfiner (1975) and Delhomme (1976). This discussion is a summary of Chapter 2 of Hughes and Lettenmaier (1980).

Matheron (1971) defines a regionalized variable (ReV) as a function, $Z(\underline{x})$, which gives the value of the phenomena under study at a point $\underline{x} = (x,y)$ in the plane. A ReV is analogous to an observation z_i which is the realization of a random variable, Z . In this case the ReV is the realization of a random function; the ReV is a set of measurements, not just a single measurement. Typical examples of ReV's are measured hydraulic conductivity, transmissivity, potentiometric, and contaminant plume data.

Definition: A random function $Z(\underline{x})$ is said to be strictly stationary if the joint probability density function (f) of m arbitrary points is invariant under translation (Delhomme 1976), i.e.,

$$f[Z(\underline{x}_1), \dots, Z(\underline{x}_m)] = f[Z(\underline{x}_1 + \underline{h}), \dots, Z(\underline{x}_m + \underline{h})]$$

Definition: A random function $Y(\underline{x})$ is said to be weakly stationary (or second order stationary) if only its first two moments are invariant under translation,

$$E[Z(\underline{x})] = \text{constant} \quad \text{for all } \underline{x}$$

and

$$\text{Cov}[Z(\underline{x}_1), Z(\underline{x}_2)] = C(\underline{x}_1 - \underline{x}_2)$$

i.e., the covariance depends only on the difference between the points $\underline{x}_1$ and $\underline{x}_2$, not the actual points themselves. The covariance is said to be isotropic if it depends only on the distance between two points, i.e.,

$$\text{Cov} [Z(\underline{x}_1), Z(\underline{x}_2)] = C(|\underline{x}_1 - \underline{x}_2|)$$

Weak stationarity can still be too restrictive since one must know $E[Z(\underline{x})]$ (even if it is a constant), and the variance of $Z(\underline{x})$ must be finite. The result is that most of the work on kriging originated around the following concept, called the intrinsic hypothesis.

Definition: A random function $Z(\underline{x})$ satisfies the intrinsic hypothesis if the differences, $Z(\underline{x} + \underline{h}) - Z(\underline{x})$, have first and second moments which depend only on $\underline{h}$. The semivariogram (or variogram), γ , is defined as (Matheron 1971)

$$\begin{aligned}\gamma(\underline{h}) &= \frac{1}{2} E [[Z(\underline{x} + \underline{h}) - Z(\underline{x})] - E[Z(\underline{x} + \underline{h}) - Z(\underline{x})]]^2 \\ &= \frac{1}{2} \sigma^2 [Z(\underline{x} + \underline{h}) - Z(\underline{x})]\end{aligned}$$

The following properties hold for $\gamma(\underline{h})$:

1. $\gamma(\underline{h}) = \gamma(-\underline{h})$
2. $\gamma(0) = 0$

(Note: it is not necessary that $\lim_{h \rightarrow 0} \gamma(\underline{h}) = 0$ as $h \rightarrow 0$)

3. $\lim_{h \rightarrow \infty} \gamma(\underline{h}) = \text{Cov}(0)$ [if $\text{Cov}(0)$ exists]

An important case arises when $E[Z(\underline{x})] = \text{constant}$. The semivariogram may be reduced to

$$\gamma(\underline{h}) = \frac{1}{2} E [Z(\underline{x} + \underline{h}) - Z(\underline{x})]^2 \quad (1)$$

and an estimate is given by (Delfiner 1975)

$$\gamma(\underline{h}) = \frac{1}{2N(\underline{h})} \sum_{i=1}^{N(\underline{h})} [Z(\underline{x}_i + \underline{h}) - Z(\underline{x}_i)]^2 \quad (2)$$

where $N(\underline{h})$ = number of pairs of observations separated by $\underline{h}$, and $Z(\underline{x}_i + \underline{h})$ and $Z(\underline{x}_i)$ are observations. When applying the above formula, it is common practice to assume isotropy (i.e., $\underline{h} = |\underline{h}| = h$), since the amount of data is generally not sufficient to model several variograms in different directions. Isotropy will be assumed throughout the rest of this discussion.

Some commonly used models of the variogram are shown in Figure 2.1.

One final point regarding the variogram should be emphasized. The theoretical restriction that $\gamma(0) = 0$ does not preclude $\lim_{h \rightarrow 0} \gamma(h) \neq 0$. An illustration of this may be seen in Figure 2.2. The difference $\lim_{h \rightarrow 0} \gamma(h) - \gamma(0) = C$ is called the nugget effect. This may arise in two situations:

1. In mining assays (very discontinuous and accurate data), the presence of veins of ore means that one pair of measurements (both inside or both outside the vein) may result in a small difference $[Z(\underline{x} + \underline{h}) - Z(\underline{x})]$ while another pair of measurements separated by the same distance (but this time with one inside and one outside the vein) may produce a large difference. As a result, the variance of the differences is large, even at small distances.
2. Where the data values may be inaccurate due to instrumentation, C can represent measurement error. C is a good indication of the degree of discontinuity of the random function. The larger C is, the more discontinuous the random function (Delhomme 1976). The magnitude of C is also related to predictability; if C is very large, any given observation contains very little information about the process at neighboring points and the process will be difficult to predict, regardless of the number of observations.

2.1 DRIFT

The drift is the structural or deterministic component of the random function. From a statistical viewpoint, the drift is the expected value of the random function. While the variogram attempts to describe the random fluctuations of the random function, the drift represents the slowly varying, continuous features. Therefore, the drift may be approximated within a restricted neighborhood by an analytical expression of the following form:

a) POWER FUNCTION ($\gamma(h) = w|h|^\lambda$)

b) SPHERICAL ($\gamma(h) = w \left[\frac{3|h|}{2a} - \frac{1}{2} \left(\frac{|h|}{a} \right)^2 \right]$
 $h \leq a$;
 $= w \quad h > a$)

c) EXPONENTIAL ($\gamma(h) = w \left[1 - e^{-\frac{|h|}{a}} \right]$)

c) GAUSSIAN ($\gamma(h) = w \left[1 - e^{-\left(\frac{|h|}{a} \right)^2} \right]$)

FIGURE 2.1. Common Variogram Models for Scaling Factor $w = 1.0$
 (from Delhomme 1976)

FIGURE 2.2. Linear Variogram with Nugget Effect, C

$$m(\underline{x}) = \sum_{i=1}^L a_i f_i(\underline{x})$$

where $m(\underline{x})$ is the drift at a point $\underline{x}$, the a_i are constant coefficients and the f_i are arbitrary functions of the point $\underline{x}$. In the plane one might let $m(\underline{x}) = a_1 + a_2x + a_3y$ or $m(\underline{x}) = a_1 + a_2\cos(x) + a_3\sin(y)$. When $Cov(h)$ is known, the a_i may be determined using linear regression analysis (Olea 1975). When $\gamma(h)$ is known, the drift can only be determined up to an arbitrary constant, i.e.,

$$m(\underline{x}) = \sum_{i=1}^L a_i f_i(\underline{x})$$

where a_1 cannot be determined (Olea 1975). One of the primary difficulties in the application of the theory of Rev's should now become apparent. The variogram must be known so that the drift may be determined. However, knowledge of the drift must be assumed in estimating the variogram (unless, of course, the drift is constant). In practice, of course, neither is usually known and both must be estimated from the data.

2.2 GENERALIZED INCREMENTS AND GENERALIZED COVARIANCES

A way to statistically quantify the randomness of a regionalized variable that eliminates the need to estimate the drift altogether has been developed by Delfiner (1975). The key concept is generalized increments, which is analogous to the differencing process used in time series analysis. Consider the case in which the drift is constant. Then the 1st order difference (0th order increment) of the observations $Z(\underline{x} + \underline{h}) - Z(\underline{x})$ has eliminated the drift. Notice, however, that any sum

$$\sum_{i=1}^n \lambda_i Z(\underline{x}_i)$$

for which

$$\sum_{i=1}^n \lambda_i = 0$$

will also eliminate the constant drift. Now consider a random function $Z(\underline{x})$ with a linear drift. In this case, the 2nd order difference (1st order increment) $Z(\underline{x} + \underline{h}) - 2Z(\underline{x}) + Z(\underline{x} - \underline{h})$ eliminates the drift. Similarly, any sum

$$\sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad \text{for which} \quad \sum_{i=1}^n \lambda_i = 0 \quad , \quad \sum_{i=1}^n \lambda_i x_i = 0$$

and

$$\sum_{i=1}^n \lambda_i y_i = 0 \quad (\underline{x}_i = (x_i, y_i))$$

will also eliminate a linear drift. In general a generalized increment of order k is defined as follows:

Definition: A linear combination of observations,

$$\sum_{i=1}^n \lambda_i Z(\underline{x}_i)$$

is a generalized increment of order k if and only if

$$\sum_{i=1}^n \lambda_i f(\underline{x}_i) = 0$$

where $f(\underline{x})$ is a polynomial of order less than or equal to k . In the plane, this means

$$k = 0 \quad \sum_{i=1}^n \lambda_i = 0$$

$$k = 1 \quad \sum_{i=1}^n \lambda_i = 0, \quad \sum_{i=1}^n \lambda_i x_i = 0, \quad \sum_{i=1}^n \lambda_i y_i = 0,$$

$$k = 2 \quad \sum_{i=1}^n \lambda_i = 0, \quad \sum_{i=1}^n \lambda_i x_i = 0, \quad \sum_{i=1}^n \lambda_i y_i = 0, \quad \sum_{i=1}^n \lambda_i x_i^2 = 0$$

$$\sum_{i=1}^n \lambda_i y_i^2 = 0, \quad \sum_{i=1}^n \lambda_i x_i y_i = 0$$

Clearly, an increment of order 0 filters out a constant drift, and increment of order 1 filters out a linear drift, and an increment of order 2 filters out a quadratic drift.

Definition: $Z(\underline{x})$ is known as an intrinsic random function of order k if there exists a generalized covariance function, K , such that for any k th order generalized increment

$$\sum_{i=1}^n \lambda_i Z(\underline{x}_i)$$

$$\text{Var} \left[\sum_{i=1}^n \lambda_i Z(\underline{x}_i) \right] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(\underline{x}_i - \underline{x}_j) \quad (3)$$

Matheron (1973) has shown that any isotropic generalized covariance actually defines an entire family of functions which are all equivalent up to an arbitrary even-powered polynomial of order $\leq 2k$. That is, a generalized covariance $K(\underline{h})$ actually defines a family of functions F where $K^*(\underline{h}) \in F$ if and only if $K(\underline{h}) - K^*(\underline{h})$ is an even polynomial of degree $\leq 2k$. For example, the two functions $K(\underline{h}) = -5h + 2h^3$ and $K^*(\underline{h}) = -5h + 4h^2 + 2h^3$ are members of the same family and, therefore, equivalent ($h = |\underline{h}|$). For this reason, only the essential odd powers are used when writing a generalized covariance. The only isotropic polynomial generalized covariances are given by:

Drift	Order (k)	$K(\underline{h})$
Constant	0	$K(\underline{h}) = C\delta + a_0h$
Linear	1	$K(\underline{h}) = C\delta + a_0h + a_1h^3$
Quadratic	2	$K(\underline{h}) = C\delta + a_0h + a_1h^3 + a_2h^5$

where

$$C > 0, \quad a_0 < 0, \quad a_1 > -\frac{10\sqrt{a_0a_2}}{3}, \quad a_2 < 0$$

and δ is the Dirac-delta function

$$\begin{aligned} \delta &= 1 & \text{if } h &= 0 \\ \delta &= 0 & \text{if } h &\neq 0. \end{aligned}$$

The constraints on the coefficients C , a_0 , a_1 , and a_2 are necessary to ensure that the variance of Equation (3) is nonnegative. Polynomial generalized covariances with coefficients satisfying these constraints are said to be "admissible."

2.3 KRIGING

At this point, let us reconsider the initial objective. Our objective is to estimate the value of a regionalized variable Z at a point $\underline{x}_0$. The estimate of $Z(\underline{x}_0)$, denoted by $Z^*(\underline{x}_0)$, is expressed as a linear combination of the observed data

$$Z^*(\underline{x}_0) = \sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad (4)$$

The objective is to find weights λ_i (kriging weights) such that

$$E(Z_0^* - Z_0) = 0 \quad (\text{unbiasedness})$$

$$\text{Var}(Z_0^* - Z_0) = E(Z_0^* - Z_0)^2 \quad \text{is minimized (minimum variance).}$$

For conciseness of notation we have written $Z_i = Z(\underline{x}_i)$ for $i = 0, \dots, n$. Using Equation (4), the unbiasedness condition is rewritten as

$$E \left\{ \left[\sum_{i=1}^n \lambda_i Z_i \right] - (Z_0) \right\} = 0$$

$$\left[\sum_{i=1}^n \lambda_i E(Z_i) \right] - E(Z_0) = 0 \quad .$$

Using the previous definition of a polynomial drift of order less than or equal to k ,

$$m(\underline{x}) = \sum_{j=1}^L a_j f_j(\underline{x})$$

where the collection of functions, $\{f_j(\underline{x})\}$, is given by $\{1\}$ for $k = 0$; $\{1, x, y\}$ for $k = 1$; $\{1, x, y, x^2, y^2, xy\}$ for $k = 2$; etc. The unbiased condition becomes

$$\left\{ \sum_{i=1}^n \lambda_i \sum_{j=1}^L a_j f_j(\underline{x}_i) \right\} - \sum_{j=1}^L a_j f_j(\underline{x}_0) = 0 \quad .$$

Reversing the order of summation, the equation becomes

$$\sum_{j=1}^L a_j \left\{ \sum_{i=1}^n \lambda_i f_j(\underline{x}_i) - f_j(\underline{x}_0) \right\} = 0$$

Since this latter expression must hold for all possible values a_j , it must be true that

$$\sum_{i=1}^n \lambda_i f_j(\underline{x}_i) = f_j(\underline{x}_0) \quad , \text{ for } j = 1, \dots, L$$

For example, in the case of linear drift ($k = 1$) the constraints are given by

$$\sum_{i=1}^n \lambda_i = 1 \quad , \quad \sum_{i=1}^n \lambda_i x_i = x_0 \quad , \quad \sum_{i=1}^n \lambda_i y_i = y_0$$

The following theorem gives an expression for $\text{Var}(Z_0^* - Z_0)$ (i.e., the estimation variance) in terms of the λ_i and the generalized covariance $K(h)$ for Z an intrinsic random function of order k .

Theorem: Let $Z(\underline{x})$ be an intrinsic random function of order k . If

$$Z^*(\underline{x}_0) = \sum_{i=1}^n \lambda_i Z(\underline{x}_i)$$

is an unbiased estimate of $Z(\underline{x}_0)$ then

$$\begin{aligned} \text{Var}(Z(\underline{x}_0) - Z^*(\underline{x}_0)) &= K(0) - 2 \sum_{i=1}^n \lambda_i K(\underline{x}_i - \underline{x}_0) \\ &\quad + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(\underline{x}_i - \underline{x}_j) \end{aligned} \quad (5)$$

This theorem becomes obvious by observing that $Z(\underline{x}_0) - Z^*(\underline{x}_0)$ is a generalized increment of order k .

The kriging estimation problem involves finding the values of the kriging weights λ_i , $i = 1, \dots, n$, such that Equation (5) is minimized subject to the constraints

$$\sum_{i=1}^n \lambda_i f_j(\underline{x}_i) = f_j(\underline{x}_0) \quad , \quad \text{for } j = 1, \dots, L$$

The resulting system of equations, known as the kriging system, is as follows:

$$\begin{aligned} \sum_{j=1}^n \lambda_j K(\underline{x}_i - \underline{x}_j) + \sum_{s=1}^L \mu_s f_s(\underline{x}_i) &= K(\underline{x}_i - \underline{x}_0) \quad i = 1, \dots, n \\ \sum_{i=1}^n \lambda_j f_s(\underline{x}_j) &= f_s(\underline{x}_0) \quad s = 1, \dots, L \end{aligned} \quad (6)$$

or, in matrix terms

$$CW = P$$

where

$$C = \begin{bmatrix} K_{11} & \dots & K_{1n} & 1 & f_1(\underline{x}_1) & \dots & f_L(\underline{x}_1) \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ K_{n1} & \dots & K_{nn} & 1 & f_1(\underline{x}_n) & \dots & f_L(\underline{x}_n) \\ 1 & \dots & 1 & 0 & \dots & \dots & 0 \\ f_1(\underline{x}_1) & \dots & f_1(\underline{x}_n) & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ f_L(\underline{x}_1) & \dots & f_L(\underline{x}_n) & 0 & \cdot & & 0 \end{bmatrix} , \quad W = \begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ \mu_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_L \end{bmatrix}$$

$$\text{and } P_T = [K_{10}, \dots, K_{n0}, 1, f_1(\underline{x}_0), \dots, f_L(\underline{x}_0)]$$

Thus, to solve Equation (6) all that is needed is the order of the drift (assuming polynomials $f_1, \dots, f_L$) and the coefficients of the generalized covariance function (assuming polynomial K). Equation (6) has a unique solution for λ_i , $i = 1, \dots, n$, provided that the drift functions $f_1, \dots, f_L$ are algebraically linearly independent, i.e.,

$$\sum_{j=1}^L a_j f_j(\underline{x}) = 0 \quad \text{for all } \underline{x} \text{ if } a_j = 0 \text{ for } j = 1, \dots, L$$

Solving the kriging system [Equation (6)] is the basis of the kriging estimation routine KRIEST, and fundamental to the drift order estimation routine, FINDK, and the generalized covariance estimation routine, KHREG.

2.4 DRIFT ORDER IDENTIFICATION

In order to krig a regionalized variable $Z(\underline{x})$ using Equation (6), it is necessary to model $Z(\underline{x})$ as an intrinsic random function of order k . To determine the order of drift (k), each observation is successively deleted from the data base and estimated from the remaining observations using the generalized covariance function

$$K(h) = -h$$

and various values of k . This particular generalized covariance function is valid for any value of k . Typically, k is chosen to be 0, 1, or 2.

The optimum choice for k depends on the residuals

$$\{Z_i - Z_i^k\} \text{ where } i = 1, \dots, n \text{ and } k = 0, 1, \text{ or } 2$$

where Z_i is the i th observation and Z_i^k is the kriged estimate of Z_i derived by deleting Z_i from the data base and letting the order of the drift equal k .

Three criterion are used to compare the sets of residuals:

1. For each i the absolute values of residuals, $|Z_i - Z_i^k|$, $k = 0, 1, 2$, are ranked according to magnitude. The value of k with the smallest average rank is the preferred order of drift value. A Kruskal-Wallis test (Kruskal and Wallis 1952) may be performed to see if there is a significant difference in average ranks between the values of k .
2. For each value of k , the mean square error (MSE) is calculated,

$$MSE_k = \sum_{i=1}^n (Z_i - Z_i^k)^2 / n$$

and the minimum MSE_k is selected.

3. By definition of a generalized increment, the expected value of a residual, $E(Z_i - Z_i^k)$, should be equal to zero if the true value of k is less than or equal to the assumed value. Therefore, for each value of k compute

$$Z_k = \bar{e}_k / s(\bar{e}_k)$$

where

$$\bar{e}_k = \sum_{i=1}^n (Z_i - Z_i^k) / n$$

and

$$s(\bar{e}_k) = \left[\sum_{i=1}^n (Z_i - Z_i^k - \bar{e}_k)^2 / (n - 1) \right]^{1/2}$$

Under the hypothesis that the true value of the order of drift is less than or equal to the assumed value k , Z_k is asymptotically a standard normal random variable. Thus, if $|Z_k|$ is greater than 1.96, it is 95% probable that the true order of drift is greater than k . This test also helps to decide whether any

order of polynomial drift is adequate to model the regionalized variable. In most cases, however, $k = 0, 1$ or 2 is adequate.

These three procedures are implemented in the routine FINDK. It is our experience that the value of k with minimum mean square error and/or average rank is selected; the hypothesis test of Criterion 3 is most appropriate when the sample size of the data base is large (>50).

2.5 GENERALIZED COVARIANCE IDENTIFICATION

Once k is chosen, the coefficients for polynomial generalized covariance are estimated as follows: Let

$$Z_r(\lambda) = \left[\sum_{j=0}^n \lambda_j^r Z(\underline{x}_j) \right]$$

be a generalized increment of order k that is the kriging error corresponding to deleting the r th value from the data base and then kriging an estimate. By Equation (3),

$$\text{Var}(Z_r(\lambda)) = E(Z_r(\lambda))^2 = \sum_{i=0}^n \sum_{j=0}^n \lambda_i^{(r)} \lambda_j^{(r)} K(\underline{x}_i - \underline{x}_j)$$

Since the model for a polynomial generalized covariance of order k is

$$K(h) = C\delta + \sum_{p=0}^k a_p |h|^{2p+1}$$

substitution yields

$$E(Z_r(\lambda))^2 = C \sum_{i=0}^n (\lambda_i^{(r)})^2 + \sum_{p=1}^k a_p \sum_{i=0}^n \sum_{j=0}^n \lambda_i^{(r)} \lambda_j^{(r)} |\underline{x}_i - \underline{x}_j|^{2p+1} \quad (7)$$

Letting

$$T_0^r = \left(\sum_{i=0}^n \lambda_i^{(r)} \right)^2$$

$$T_{2p+1}^r = \sum_{i=0}^n \sum_{j=0}^n \lambda_i^{(r)} \lambda_j^{(r)} |\underline{x}_i - \underline{x}_j|^{2p+1}$$

Equation (7) can be rewritten as

$$E(Z_r(\lambda))^2 = CT_0^r + \sum_{p=1}^k a_p T_{2p+1}^r \quad (8)$$

which is linear in the unknown parameters. Delfiner (1976) recommends regressing $E[Z_r(\lambda)]^2$ upon T_0^r, T_{2p+1}^r $p = 1, \dots, k$ with weights $w_i^2 = (T_i)^{-2}$ to determine the coefficients. This does not appear to be essential; test runs using $w_i^2 = 1.0$ showed no significant difference. All polynomial generalized covariances of order k should be tried [e.g., for $k = 1$ try $K(h) = C\delta, C\delta + a_0h, C\delta + a_0h + a_1h^3, C\delta + a_1h^3, a_0h, a_0 + a_1h^3$]. Each regression which yields a set of admissible parameters is used to re-estimate the observations, yielding a new set of errors and T^r values. The regression [Equation (8)] is performed again and the iteration is continued until the parameters converge. We found that this convergence is not always straightforward. Sometimes the parameters will oscillate between admissible and inadmissible values. The convergence pattern may be quite complex but in the cases tested the parameters always converged (although occasionally to inadmissible values).

Once several sets of admissible parameters have been estimated it remains to choose the best set. Recall that $Z_r(\lambda)$ is a kriging error and that $E(Z_r(\lambda))^2 = \text{Var}(Z_r(\lambda))$. This is one estimate of the kriging variance. Another estimate, inherent to the kriging system, is σ_k^2 defined from Equation (5). Clearly, if the estimate of $K(h)$ is correct, these two quantities should be approximately equal. Therefore, the quantity

$$\rho = \frac{\sum_r Z_r(\lambda)^2}{\sum_r \sigma_{k,r}^2} \quad (9)$$

should be close to 1. In practice, a jackknife estimator of ρ is recommended (Delfiner 1975):

$$\rho = 2r - \frac{n_1 r_1 + n_2 r_2}{n_1 + n_2} \quad (10)$$

where

$$r = \frac{\sum_r Z_r(\lambda)^2}{\sum_r \sigma_{k,r}^2}$$

$$r_1 = \frac{\sum_{r \in I_1} Z_r(\lambda)^2}{\sum_{r \in I_1} \sigma_{k,r}^2} \quad r_2 = \frac{\sum_{r \in I_2} Z_r(\lambda)^2}{\sum_{r \in I_2} \sigma_{k,r}^2}$$

(It is sufficient to let $I_1 = \{1, \dots, n/2\}$, $I_2 = \{n/2 + 1, \dots, n\}$.)

This estimation technique is implemented in KHREG. The initial values of T_0^r and T_{2p+1}^r , $p = 1, \dots, k$ are derived from kriging weights corresponding to $C = 1$, $a_0 = -1$, $a_1 = 1$, $a_2 = -1$ (these coefficients are admissible). For example, if the form of the generalized covariance is given by $C\delta + a_0 h + a_2 h^3$, then the initial estimate of $K(h)$ is given by

$$K(h) = 1 - h + h^3$$

The optimum generalized covariance is that which has a minimum mean square error and jackknife statistic (ρ) not too far from 1.0 ($0.75 < \rho < 1.25$). For small sample sizes ($n < 50$) the jackknife estimate criterion is not as accurate, and the minimal mean square criterion for selection of $K(h)$ should be used.

2.6 REFERENCES

- Delfiner, P. 1975. "Linear Estimation of Non-Stationary Phenomena." In Advanced Geostatistics in the Mining Industry, pp. 49-68. Reidel, Dordrecht, Holland.
- Delhomme, J. P. 1976. Application of the Theory of Regionalized Variables to the Hydrologic Sciences. Dissertation, Ecole des Mines de Paris, Fontainebleau, France.
- Hughes, J. P., and D. P. Lettenmaier. 1980. "Aquatic Monitoring: Data Analysis and Network Design Using Regionalized Variable Theory." University of Washington Department of Civil Engineering, Seattle, Washington.
- Kruskal, W. H., and W. A. Wallis. 1952. "Use of Ranks in One-criterion Analysis of Variance." J. American Statistical Association 47:583-621.
- Matheron, G. 1971. "The Theory of Regionalized Variables and its Applications." Les Cahiers du CMM; fasc. 5. Fontainebleau, France.
- Matheron, G. 1973. "The Intrinsic Random Functions and Their Applications." Adv. in Applied Probability 5:439-468.

3.0 USER INSTRUCTIONS AND SAMPLE CASES

This section describes the use of the data retrieval, sorting, geostatistical and contour plotting routines. An example will be shown using all tritium concentration data (units = PCI/L) that were gathered between April and June 1967 within the rectangular subregion of the Hanford Site bounded by the following coordinates (plant coordinate system):

upper left	N52000', W47000'
upper right	N52000', E20000'
lower right	S12000', E20000'
lower left	S12000', W47000'

The step-by-step execution of this software to produce tritium plume contour plots is included below.

3.1 KRGRTV

KRGRTV retrieves all data from the CIRMIS data base management system for a user-defined data type (contaminant, hydrograph, or temperature), time period, and Hanford Site subregion. If more than one measurement is made during the time period, the average value is written to the output file. Prompts from KRGRTV and required responses from the user during program execution are listed below:

1. After logging on the system: User response = RUN KRGRTV
2. Prompt = ENTER FILE-Q DISK UNIT I.E. DB2, DB1 OR DBØ . User response = DB1 . (This is the disk unit in which the CIRMIS data base is stored.)
3. Prompt = ENTER OUTPUT FILE NAME . User response = TØ667.DAT. (KRGRTV will write data values to this file.)
4. Prompt = ENTER WELL NAME INPUT FILE NAME (30A1) . User response = DB1: WLHØRØ383.SEQ . (This is the particular CIRMIS data base to be searched.)

5. Prompt = DO YOU NEED A NEW DATA TYPE? (Y or N) . User response = Y . (Enter Y for yes if this is the initial search, enter N for no if this is not the initial search and you do not want to change data type.)
6. Prompt = ENTER DATA TYPE, 1 = CTM, 2 = HYD, 3 = TMP (I2) . User response = 1 . (Enter data type 1 = CTM = Contaminant Concentration, 2 = HYD = hydrograph, 3 = TMP = temperature; this entry is skipped if user response to 5. is N.)
7. Prompt = ENTER CONTAMINANT NUMBER . User response = 3 (3 corresponds to tritium). (Enter contaminant number according to Table 1. This entry is skipped unless user response to 6. is 1 (CTM).)
8. Prompt = DO YOU NEED A NEW TIME SPAN? (Y or N) . User response = Y . (Enter Y for yes if initial run of KRGRTV or if you wish to use another time period, N for no otherwise.)
9. Prompt = SPECIFY START MONTH AND YEAR 2(I2) . User response = 0467 . (Enter month and year in 2I2 format, e.g., April, 1967 corresponds to 0467.)
10. Prompt = SPECIFY END MONTH AND YEAR (2I2) . User response = 0667 . (Enter month and year 2I2 format as in 9. Entries 9. and 10. are skipped if the response to 8. is N.)
11. Prompt = DO YOU WANT ALL WELLS? (Y or N) . User response = N . (Enter Y (yes) or N (no).)
12. Prompt = DO YOU NEED TO DEFINE A NEW BOX? (Y or N) . User response = Y . (Enter Y for yes or N for no if a new rectangular subregion of the Hanford Reservation is to be searched. For the initial run enter Y.)
13. Prompt = Using PLANT or RICHLAND COORDINATE SYSTEM . User response = PLANT . (CIRMIS utilizes a local coordinate system with origin on the Hanford Reservation, denoted the PLANT System or the city of RICHLAND COORDINATE SYSTEM. For most work the PLANT coordinate system is most adequate.)
- 14-17. Prompts = ENTER Y, X COORDINATE 1 (2, 3, or 4) TO DEFINE THE UPPER LEFT (UPPER RIGHT, LOWER RIGHT, LOWER LEFT) CORNER OF BOX AROUND REGION [2(A1, F10.0)] . User response = N52000., W47000., (N52000., E20000.; S12000., E20000.; S12000., W47000.) . Steps 13-17. are omitted if the user response to 12. is N(no). (Enter the coordinates of the corners of the rectangular subregion to be searched.)

KRGRTV searches DB1: WLHDR083.SEQ for the proper type of data in the indicated subregion and the particular time period and writes it to the formatted file T0667.DAT.

18. Prompts = ARE YOU FINISHED? (Y or N) . User response = N . (Enter Y for yes if finished or N for no if you wish to retrieve more data to OUTPUT.FIL. In the latter case you return to Step 5.)

KRGRTV echoes all input to the lineprinter and writes the indicated CIRMIS data volume to T0667.DAT. Figure 3.1 is a sample listing of T0667.DAT for the tritium plume test case.

```

COMPARISON OF INTEREST          FREQUENCY (PCI/L) -- DATA FROM 467 TO 667
46
108275.0    38400.0  0.320E+04
108275.0    38400.0  0.250E+04
111000.0    50500.0  0.180E+04
111000.0    50500.0  0.230E+04
121201.0    56300.0  0.240E+04
121201.0    56300.0  0.240E+04
124000.0    63200.0  0.270E+04
124211.0    62500.0  0.240E+04
123211.0    62500.0  0.250E+04
112781.0    62400.0  0.300E+04
121000.0    60000.0  0.230E+04
105875.0    70900.0  0.215E+04
121000.0    79000.0  0.265E+04
121000.0    79000.0  0.280E+04
114507.0    82711.0  0.320E+04
114500.0    82869.0  0.240E+04
71622.0     83210.0  0.250E+04
94037.0     83815.0  0.130E+03
82630.0     84602.0  0.530E+06
61682.0     86333.0  0.250E+04
65850.0     87825.0  0.320E+04
50740.0     87953.0  0.320E+04
113760.0    89148.0  0.270E+04
121017.0    89304.0  0.240E+04
121017.0    89304.0  0.240E+04
88610.0     89389.0  0.540E+06
75685.0     92809.0  0.640E+06
94446.0     94665.0  0.140E+07
100719.0    96027.0  0.380E+06
68703.0     96696.0  0.218E+05
78322.0     99507.0  0.250E+07
78322.0     99507.0  0.550E+05
78322.0     99507.0  0.210E+05
78322.0     99507.0  0.210E+07
99825.0     103700.0  0.320E+04
106430.0    108849.0  0.210E+04
75574.0     109404.0  0.210E+04
86124.0     109692.0  0.665E+06
97275.0     111474.0  0.150E+05
66937.0     110994.0  0.260E+07
66901.0     114274.0  0.460E+06
88371.0     115093.0  0.320E+04
74341.0     115450.0  0.210E+04
80650.0     119000.0  0.250E+04
67071.0     119206.0  0.350E+04
55733.0     118840.0  0.530E+04

```

FIGURE 3.1. Sample Listing of KRGRTV Output (T0667.DAT)

TABLE 3.1. Contaminant Code for the CIRMIS
Data Base Management System

1	--	TOTAL ALPHA
2	--	TOTAL BETA
3	--	TRITIUM
4	--	NITRATE
5	--	STRONTIUM-90
6	--	CESIUM-137
7	--	COBALT-60
8	--	URANIUM-238
9	--	PLUTONIUM-239
10	--	RUTHENIUM-106
11	--	CHROMIUM
12	--	FLUORIDE
13	--	MAGNESIUM
14	--	IRON
15	--	PHOSPHATE
16	--	CHLORIDE
17	--	COPPER
18	--	HARDNESS
19	--	SOLIDS
20	--	NITRATE (ION)
21	--	TOTAL ORGANIC CARBON
22	--	PH
23	--	SULFATE
24	--	SODIUM
25	--	CALCIUM
26	--	TOTAL GAMMA
27	--	BICARBONATE ION
28	--	CARBONATE ION
29	--	TOTAL POTASSIUM
30	--	SPECIFIC CONDUCTIVITY
31	--	BORON
32	--	LOW ALPHA

3.2 KRIGTRAN

KRIGTRAN transforms the output data file of KRGRTV into a form compatible with the geostatistical routines FINDK, KHREG, and KRIEST. KRIGTRAN reads the title and number of values from the KRGRTV output file, scales the x-location, y-location, and parameter value (contaminant, hydrograph, and temperature) by a factor of 1000, and screens the data base for replicate measurement values at well locations. In the case of replicate measurement values, KRIGTRAN uses the maximum value at a well location. Finally, KRIGTRAN writes the scaled values, the title and number of data values to the output file.

The following steps are used to execute KRIGTRAN on the Geosciences Research and Engineering PDP 11/70 computer:

1. After logging on the system: User response = RUN KRIGTRAN
2. Prompt = ENTER CIRMIS FILE NAME . User response = T0667.DAT . (The user enters the output file of KRGRTV.)
3. Prompt = ENTER THE OUTPUT FILE NAME . User response = T0667.TMP . (The user enters the name of the data file to be used as input to the geostatistical routines.)

A listing of the KRIGTRAN output file (T0667.TMP) created using T0667.DAT as input is shown in Figure 3.2. Notice that the x and y coordinates and the tritium concentration values are scaled by 1000. The resulting tritium concentration is in units of PCI/ML. T0667.TMP will now be used as input to FINDK, KHREG, and KRIEST.

108.2750	38.4000	0.3200E+01
111.0000	50.5000	0.2500E+01
121.2610	56.3000	0.2400E+01
124.0000	63.7000	0.2700E+01
123.2110	62.5000	0.2500E+01
112.7810	62.9000	0.3000E+01
121.0000	66.0000	0.2300E+01
105.0750	70.9000	0.2450E+01
121.0000	75.0000	0.2800E+01
119.0070	82.7110	0.3200E+01
114.5000	82.8690	0.2400E+01
71.0220	83.2100	0.2500E+01
94.0370	83.8150	0.1300E+00
82.6390	84.0020	0.5300E+03
61.6820	86.3330	0.2500E+01
65.8500	87.8250	0.3200E+01
50.7400	87.9530	0.3200E+01
113.7800	89.1000	0.2700E+01
121.0070	89.3000	0.2400E+01
68.6100	89.3890	0.5800E+03
75.6850	92.8090	0.6400E+03
94.4460	94.8650	0.1400E+04
100.7130	96.0270	0.3800E+03
68.7030	96.8960	0.2180E+02
78.3220	99.5070	0.2500E+04
99.8250	103.7000	0.3200E+01
108.4300	106.6490	0.2100E+01
75.5740	109.4040	0.2100E+01
86.1230	109.6920	0.6650E+03
97.2750	111.4740	0.1500E+02
66.9370	110.9940	0.2600E+04
60.9010	114.2740	0.4600E+03
86.3710	115.0930	0.3200E+01
74.3410	115.8600	0.2100E+01
50.6500	119.0000	0.2500E+01
67.0910	119.2060	0.3500E+01
55.7330	119.8400	0.5300E+01

CONTINUED AT DE INTEREST TRITIUM (PCI/L) -- DATA FROM 467 TO 667
 37 SORTED VALUES
 X AND Y POSITION UNITS=1000FT,CUNC=PCI/ML
 THESE ARE THE TRANSFORMED UNITS FOR KRIGING

FIGURE 3.2. Sample Listing of KRIGTRAN Output (T0667.TMP)

Given a data base, FINDK will calculate the average rank, the mean square error, and a normalized error statistic for the order of drift $k = 0, 1$, and 2. FINDK calculates these statistics by successively deleting each data point from the data base and then kriging the observation using the generalized covariance function

Refer to Chapter 2.0 for details on identification of drift order.

1. After logging on the system: User response = RUN FINDK .
2. Prompt = ENTER THE NUMBER OF DATA VALUES . User response = N . (The user must enter the number (N) of ordered triples in the observed data base disk file; $N < 100$. An ordered triple is (X,Y,Z) where (X,Y) is the spatial position and Z is the observed value.)
3. Prompt = ENTER THE NAME OF THE INPUT FILE . User response = INPUTFILE.DAT . (The name of the observed data base file should be not be more than 27 characters in length. Each of the N records of the input file contains (X,Y,Z) with a * type format; for example,

```
-----2.50-----3.50----457.63      where - indicates
                                     a space.)
```

- ### 3.7

FINDK echoes the input parameters and data to the lineprinter and calculates for each value of k (0, 1, and 2) the observed values, estimated values, residuals (estimated minus observed), the average rank value, the mean square error, and the normalized error statistic.

The user selects the order of drift (k) with minimum average rank and/or mean square error. The standardized error statistic should be near zero for the selected k, typically between -1.0 and 1.0.

Figure 3.3 shows a sample run of FINDK using a set of tritium plume data (T0667.TMP) as input.

```
***** FINDK --A COMPUTER PROGRAM TO ESTIMATE THE ORDER OF DRIFT  OF A REGIONALIZED VARIABLE
FOR THE PURPOSE OF KRIGING
```

```
INPUT FILE NAME=T0667.TMP
NUMBER OF OBSERVATIONS= 37
NUMBER OF POINTS IN KRIGING NBHD= 8
LOG TRANSFORMATION OF Z DATA?(1=YES,0=NO) 1
DATA SET
```

X	Y	Z
108.2750	34.4000	3.2000
111.0000	50.5000	2.5000
121.2010	54.3080	2.4000
124.0000	63.2000	2.7000
123.2110	62.5650	2.5000
112.7810	62.9000	3.0000
123.0000	66.0000	2.3000
105.6750	70.9000	2.4500
121.0000	79.0000	2.8000
114.5070	82.7110	3.2000
114.5000	82.8690	2.4000
71.0220	83.2100	2.5000
94.0370	83.4150	0.1300
82.6300	84.5020	530.0000
61.6820	66.3330	2.5000
65.8500	87.8250	3.2000
50.7400	87.9530	3.2000
113.7800	89.1480	2.7000
121.0170	89.3040	2.4000
88.6100	89.3890	580.0000
75.6450	92.8090	640.0000
94.4460	94.6050	1400.0000
100.7190	95.0270	380.0000
68.7930	96.8960	21.8000
76.3220	99.5070	2500.0000
99.8250	103.7000	3.2000
106.4300	103.8490	2.1000
75.5740	109.4240	2.1000
86.1240	109.6920	665.0000
97.2750	111.4740	15.0000
86.9370	110.3940	2600.0000
66.9010	111.2740	460.0000
88.3710	115.0930	3.2000
74.3410	115.4600	2.1000
40.6500	119.0000	2.5000
67.0910	119.2060	3.5000
55.7330	114.8400	5.3000

FIGURE 3.3. Sample Output of FINDK

```

*****DEGREE OF DRIFT= 0*****
OBSERVATION      ESTIMATE      RESIDUAL
1.1632           0.9031      -0.2601
0.9163           1.0669       0.1506
0.8755           0.9557       0.0802
0.9933           0.9029      -0.0903
0.9163           0.9023       0.0460
1.0986           0.8594      -0.2392
0.8329           1.0066       0.1737
0.8961           0.2410      -0.6551
1.0296           0.9860      -0.0436
1.1532           0.8794      -0.2838
0.8755           1.1592       0.2838
0.9163           3.1383       2.2220
-2.0402          4.9556       6.9959
6.2729           3.1306      -3.1423
0.9163           0.8254      -0.0909
1.1632           1.4650       0.3019
1.1632           1.2769       0.1138
0.9933           1.7333       0.7401
0.8755           0.6360      -0.2394
6.3530           4.9176      -1.4452
6.8615           5.2952      -1.1663
7.2442           5.0444      -2.1599
5.9402           3.1853      -2.7548
3.0819           5.0757       1.9938
7.8240           4.7858      -3.0383
1.1632           4.3223       3.1592
0.7419           0.7489       0.0069
0.7419           5.2577       4.5158
6.4998           2.4818      -4.0180
2.7081           1.3265      -1.3815
7.8633           4.6958      -2.9674
6.1312           4.8208      -1.3104
1.1632           3.6961       2.5350
0.7419           1.6177       0.8758
0.9163           0.3531      -0.5632
1.2528           3.2661       2.0133
1.6677           3.8877       2.2200

```

```

*****DEGREE OF DRIFT= 1*****
OBSERVATION      ESTIMATE      RESIDUAL
1.1632           0.8579      -0.3053
0.9163           1.0662       0.1499
0.8755           0.9629       0.0874
0.9933           0.9008      -0.0924
0.9163           0.9623       0.0460
1.0986           0.8581      -0.2405
0.8329           1.0061       0.1732
0.8961           0.1307      -0.7654
1.0296           0.9866      -0.0410
1.1632           0.8815      -0.2816
0.8755           1.1596       0.2841
0.9163           2.7844       1.8681
-2.0402          4.9319       6.9721
6.2729           2.4274      -3.8454
0.9163           0.6397      -0.2766
1.1632           1.5139       0.3507
1.1632          -1.9076      -3.1308
0.9933           1.7217       0.7285
0.8755           0.0246      -0.8509
6.3530           5.0056      -1.3574

```

FIGURE 3.3. (contd)

6.4615	5.2596	-1.2019
7.2442	5.1709	-2.0733
5.9402	3.1619	-2.7783
3.0819	5.0978	2.0159
7.8240	4.7763	-3.0478
1.1032	4.4052	3.2421
0.7419	-1.8398	-2.5817
0.7419	5.3516	4.6097
6.4998	2.5892	-3.9105
2.7081	0.9152	-1.7929
7.8633	4.9093	-2.9539
6.1312	4.8704	-1.2608
1.1032	3.4237	2.2606
0.7419	1.5324	0.7904
0.9163	-1.3804	-2.2966
1.2528	2.2464	0.9936
1.6677	3.2485	1.5807

*****DEGREE OF FREEDOM= 2*****

OBSERVATION	ESTIMATE	RESIDUAL
1.1032	0.7939	-0.3692
0.9163	1.1373	0.2211
0.8755	0.6960	-0.1795
0.9933	0.8677	-0.1255
0.9163	0.9642	0.0479
1.0780	0.6435	-0.4551
0.8329	1.2764	0.4435
0.8961	0.5830	-0.3131
1.0296	1.1070	0.0774
1.1032	0.8671	-0.2961
0.8755	1.1556	0.2802
0.9163	3.5909	2.6746
-2.0402	5.6435	7.6837
6.2729	3.1715	-3.1014
0.9163	0.1215	-0.7948
1.1032	1.5869	0.4238
1.1032	18.1848	17.0216
0.9933	1.8380	0.8448
0.8755	-0.3209	-1.1964
6.3530	7.8265	1.4635
6.4615	6.4486	-0.0129
7.2442	6.6702	-0.5741
5.9402	2.3399	-3.6003
3.0819	5.3064	2.2245
7.8240	3.4043	-4.4192
1.1032	5.8616	4.6984
0.7419	2.5395	1.7976
0.7419	5.9471	5.2052
6.4998	-1.0091	-7.5089
2.7081	1.4365	-1.2715
7.8633	4.9248	-2.9384
6.1312	5.0784	-1.0528
1.1032	4.3889	3.2258
0.7419	0.6894	-0.0525
0.9163	-4.7342	-5.6505
1.2528	1.7489	0.4961
1.6677	33.8568	32.1911

OBSERVED MEAN SQUARE ERROR=	DEGREE=0	DEGREE=1	DEGREE=2
STANDARDIZED ERROR T STAT=	4.5840	5.1220	43.4681
AVERAGE RANK VALUE =	0.1954	-0.6437	1.1809
	1.7027	1.8649	2.4324

FIGURE 3.3. (contd)

On the basis of the minimum mean square error (4.5840), minimum average rank value (1.7027), and small standardized error statistic (0.1954), the choice of drift order is $k = 0$, i.e., a constant drift.

3.4 KHREG

Given a data base and order of drift (k), KHREG will calculate the coefficients of all admissible polynomial generalized covariance functions. KHREG calculates these coefficients by performing a linear regression of the squared residuals upon functions of the kriging weights. Refer to Chapter 2.0 for details of generalized covariance function identification.

Use the following steps to execute KHREG on the Geosciences Research and Engineering 11/70 computer system.

1. After logging in: User response = RUN KHREG .
2. Prompt = ENTER THE NUMBER OF DATA VALUES . User response = N . (The user must enter the number (N) of ordered triples in the observed data base disk file; $N < 100$. An ordered triple is (X,Y,Z) where (X,Y) is the spatial position and Z is the observed value.)
3. Prompt = ENTER THE NAME OF THE INPUT FILE . User response = INPUTFILE.DAT . (The user must enter the name of the observed data base file; no more than 27 characters. Each of the N records of the input file contains (X,Y,Z) with a * type format; for example,
2.50, 3.50, 457.63
or
-----2.50-----3.50----457.63 where - indicates a space.)
4. Prompt = ENTER THE NUMBER OF POINTS IN KRIGING NBHD . User response = IDIM . (The user must enter the number of data values (IDIM) to be used in the sum $\sum \lambda_i Z_i$. KHREG locates the IDIM nearest points to the position to be estimated; $IDIM < 16$.)

5. Prompt = DO YOU WISH A LOG TRANSFORM OF Z? ENTER 1 = YES, 0 = NO .
User response = ILOG . (The user must enter a value of ILOG, 1 or 0 accordingly. If ILOG equals one then KHREG analyzes $\ln(Z)$ data. The log transformation can result in a less variable data base and ensures positive predictions for the kriged surface. This latter property is useful for contaminant plume data where negative values are unrealistic.)
6. Prompt = ENTER THE ORDER OF DRIFT . User response = k . (The user must enter the order of drift (k) which corresponds to the degree of polynomial that models the slowly varying trend in the data. Typically, FINDK is run to determine k; k = 0, 1, or 2.)

KHREG echoes the input parameters and data to the lineprinter and indicates for each generalized covariance model: 1) the initial and iterative coefficient values; 2) convergence or nonadmissibility of coefficient estimates; 3) if convergence, a listing of observed values, estimated values, residuals (estimated minus observed), the kriging error standard deviations (σ), and normalized residuals (residual/ σ); and 4) if convergence, the mean square error (MSE) and the jackknife statistic rho value (R). The user selects the covariance model with the minimum mean square error with a jackknife value typically between 0.75 and 1.25. If the sample size N is less than 50 the minimum mean square error criterion should be used.

Figure 3.4 shows a sample run of KHREG with the same set of tritium data (T0667.TMP) as used for the FINDK sample case with the order of drift equal to zero.

***** KHREG== A PROGRAM THAT COMPUTES THE POLYNOMIAL GENERALIZED COVARIANCE
FOR A REGIOALIZED VARIABLE*****

INPUT FILE NAME=1007.TMP
NUMBER OF OBSERVATIONS= 37
ORDER OF DRIFT= 0
NUMBER OF POINTS IN Kriging GRID= 8
LOG TRANSFORMATION OF Z DATA?(1=YES,0=NO) 1
DATA SET

X	Y	Z
108.2750	38.4000	3.2000
111.0000	50.5000	2.5000
121.2010	58.3000	2.4000
124.0000	63.2000	2.7000
123.2110	62.5000	2.5000
112.7410	62.9000	3.0000
121.0000	68.0000	2.3000
105.6750	70.9000	2.4500
121.0000	79.0000	2.8000
114.5070	82.7110	3.2000
114.5000	82.8000	2.4000
71.0420	83.2100	2.5000
94.0070	83.8150	0.1300
82.6300	84.8020	530.0000
61.6820	86.3330	2.5000
65.8500	87.8250	3.2000
50.7400	87.9530	3.2000
113.7800	89.1400	2.7000
121.0170	89.3040	2.4000
88.0100	89.3890	580.0000
75.8850	92.8090	640.0000
94.4460	94.0050	1400.0000
100.7190	96.0270	380.0000
68.7030	96.6060	21.8000
78.3220	99.5070	2500.0000
99.8250	103.7000	3.2000
108.4300	108.8490	2.1000
75.5740	109.4040	2.1000
86.1240	109.6020	665.0000
97.2750	111.4740	15.0000
66.9370	110.9940	2600.0000
66.9010	114.2740	400.0000
88.3710	115.0930	3.2000
74.3410	115.8600	2.1000
80.8500	119.0000	2.5000
87.0910	119.2000	3.5000
55.7330	118.8400	5.3000

NEW COVARIANCE MODEL FILE

	C0	H**1	H**3	H**5
INITIAL ESTIMATES	1.0000	-1.0000	0.0000	0.0000
ITERATION 1	2.5722	-0.0928	0.0000	0.0000

FIGURE 3.4. Sample Run of KHREG

CO H**1 H**3 H**5
 ITERATION 2 5.917n 0.1435 0.0000 0.0000
 *****THIS COVARIANCE IS INADMISSABLE*****

 NEW COVARIANCE MODEL FIT

CO H**1 H**3 H**5
 INITIAL ESTIMATES 1.0000 0.0000 0.0000 0.0000

CO H**1 H**3 H**5
 ITERATION 1 5.7323 0.0000 0.0000 0.0000

CO H**1 H**3 H**5
 ITERATION 2 5.7323 0.0000 0.0000 0.0000
 ***** CONVERGENCE OF COVARIANCE COEFFICIENTS*****

OBSERVED Z	ESTIMATED Z	RESIDUAL	KRIGING SIGMA	NORMALIZED RESIDUAL
1.1632	0.9448	-0.2183	2.5395	-0.0860
0.9163	0.9757	0.0594	2.5395	0.0234
0.8755	0.9808	0.1053	2.5395	0.0415
0.9933	0.9661	-0.0272	2.5395	-0.0107
0.9163	0.9757	0.0594	2.5395	0.0234
1.0986	0.9529	-0.1457	2.5395	-0.0574
0.8329	0.9810	0.1481	2.5395	0.0583
0.4961	0.6086	-0.2875	2.5395	-0.1132
1.0296	0.9432	-0.0864	2.5395	-0.0340
1.1632	1.5677	0.4045	2.5395	0.1593
0.8755	1.6037	0.7282	2.5395	0.2867
0.9163	4.1557	3.2394	2.5395	1.2756
-2.0402	4.3808	6.4210	2.5395	2.5285
6.2729	3.8767	-2.3961	2.5395	-0.9436
0.9163	4.3433	3.4270	2.5395	1.3495
1.1632	4.1249	2.9617	2.5395	1.1663
1.1632	4.2947	3.1316	2.5395	1.2332
0.9933	2.3984	1.4052	2.5395	0.5533
0.8755	1.5591	0.6836	2.5395	0.2692
6.3630	4.2228	-2.1403	2.5395	-0.8428
6.4615	3.4099	-3.0515	2.5395	-1.2016
7.2442	4.3414	-2.9029	2.5395	-1.1431
5.9402	2.2561	-3.6841	2.5395	-1.4507
3.0819	4.0022	0.9203	2.5395	0.3624
7.5240	4.7533	-3.0708	2.5395	-1.2092
1.1632	3.9567	2.7935	2.5395	1.1001
0.7419	3.3234	2.5815	2.5395	1.0165
0.7419	4.0491	3.3071	2.5395	1.3023
6.4996	2.8128	-3.6869	2.5395	-1.4519
2.7081	3.0513	0.3433	2.5395	0.1352
7.8633	2.7947	-5.0685	2.5395	-1.9959
6.1312	3.0112	-3.1200	2.5395	-1.2286
1.1632	2.6671	1.5040	2.5395	0.5922
0.7419	4.0491	3.3071	2.5395	1.3023
0.9163	3.3878	2.4715	2.5395	0.9732
1.2528	3.2157	1.9629	2.5395	0.7730
1.0677	3.5692	1.9015	2.5395	0.7488

MSE= 6.4489 R= 1.0000

FIGURE 3.4. (contd)

 NEW COVARIANCE MODEL FIT

INITIAL ESTIMATES CO H**1 H**3 H**5
 0.0000 -1.0000 0.0000 0.0000

ITERATION 1 CO H**1 H**3 H**5
 0.0000 -0.4159 0.0000 0.0000

ITERATION 2 CO H**1 H**3 H**5
 0.0000 -0.4159 0.0000 0.0000

***** CONVERGENCE OF COVARIANCE COEFFICIENTS*****

OBSERVED Z	ESTIMATED Z	RESIDUAL	KRIGING SIGMA	NORMALIZED RESIDUAL
1.1632	0.9031	-0.2601	3.1950	-0.0814
0.9153	1.0669	0.1506	2.1769	0.0692
0.4755	0.9557	0.0802	1.6742	0.0479
0.9933	0.9029	-0.0903	0.8815	-0.1025
0.9153	0.9623	0.0460	0.8225	0.0560
1.0946	0.8594	-0.2392	1.8757	-0.1276
0.6329	1.0066	0.1737	1.5014	0.1157
0.8951	0.2410	-0.6551	2.2431	-0.2921
1.0296	0.9860	-0.0436	1.9432	-0.0225
1.1632	0.8794	-0.2838	0.3596	-0.7892
0.4755	1.1592	0.2838	0.3581	0.7925
0.9153	3.1383	2.2220	1.9262	1.1536
-2.0402	4.9556	6.9959	1.9943	3.5079
6.2729	3.1306	-3.1423	1.8897	-1.6629
0.9153	0.8254	-0.0909	1.6229	-0.0560
1.1632	1.4650	0.3019	1.4621	0.2065
1.1632	1.2769	0.1138	2.9101	0.0391
0.9933	1.7333	0.7401	1.7355	0.4264
0.6755	0.6360	-0.2394	2.0531	-0.1166
6.3630	4.9178	-1.4452	1.6650	-0.8680
6.4615	5.2952	-1.1663	1.6893	-0.6904
7.2442	5.0844	-2.1599	1.6784	-1.2868
5.9402	3.1853	-2.7548	1.7248	-1.5972
3.0619	5.0757	1.9938	1.8816	1.0596
7.8240	4.7858	-3.0383	1.7953	-1.6924
1.1632	4.3223	3.1592	1.7251	1.8313
0.7419	0.7489	0.0069	2.4108	0.0029
0.7419	5.2577	4.5158	1.6911	2.6703
6.4996	2.4816	-4.0180	1.7257	-2.3283
2.7081	1.3265	-1.3815	1.8773	-0.7359
7.6633	4.8958	-2.9674	1.4363	-2.0661
6.1312	4.8206	-1.3104	1.2612	-1.0390
1.1632	3.6981	2.5350	1.6940	1.4965
0.7419	1.6177	0.8758	1.5604	0.5613
0.9153	0.3531	-0.5632	1.8640	-0.3021
1.2526	3.2661	2.0133	1.6979	1.1858
1.6677	3.8877	2.2200	2.8469	0.7798

MSE= 4.5840 R= 1.3490

FIGURE 3.4. (contd)

The polynomial generalized covariance model to be fit in each case is indicated by the initial estimates of the coefficients. For example, the third model considered has initial coefficient estimates given by

$$C_0 = 0.0$$

$$H^{**1} = -1.0$$

$$H^{**3} = 0.0$$

$$H^{**5} = 0.0$$

which indicates that the model

$$K(h) = 0\delta + a_0h (+ 0h^3 + 0h^5)$$

is being fit. (Since the drift equals zero, model forms with cubic and quintic forms are not considered.) The final converged coefficient estimates are given by

$$C_0 = 0.0$$

$$H^{**1} = -0.4159$$

$$H^{**3} = 0.0$$

$$H^{**5} = 0.0$$

indicating the optimum generalized covariance with just a linear term is given by

$$K(h) = -0.4159h$$

The generalized covariance that appears to fit the data best (MSE = 4.5840 and R = 1.3490) is the pure linear model

$$K(h) = -0.4159h$$

The jackknife statistic ideally should be closer to 1.0, but because the mean square error is so much lower than the other models being evaluated and because the sample size is so small (N = 37), the pure linear model must be selected.

3.5 KRIEST

Given a data base, order of drift, and polynomial generalized covariance, KRIEST will krig estimates of the observed regionalized variable on a rectangular grid specified by the user. The kriged estimates may then be used as input for numerical modeling or for contour plotting. KRIEST also calculates the standard deviation of the kriging error associated with each estimate on the rectangular grid. Refer to Chapter 2.0 for details of the kriging estimation process.

Use the following steps to execute KRIEST on the Geosciences Research and Engineering PDP 11/70 computer system:

1. After logging in: User response = RUN KRIEST .
2. Prompt = ENTER THE NUMBER OF DATA VALUES . User response = N . (The user must enter the number (N) of ordered triples in the observed data base disk file; $N < 100$. An ordered triple is (X,Y,Z) where (X,Y) is the spatial position and Z is the observed value.)
3. Prompt = ENTER THE NAME OF THE INPUT FILE . User response = INPUTFILE.DAT . (The user must enter the name of the observed data base file; no more than 27 characters can be used. Each of the N records of the input file contains (X,Y,Z) with a * type format; for example,

2.50, 3.50, 457.63

or

-----2.50-----3.50----457.63

where - indicates
a space.)

4. Prompt = ENTER THE NUMBER OF POINTS IN KRIGING NBHD . User response = IDIM . (The user must enter the number of data values (IDIM) to be used in the sum $\sum \lambda_i Z_i$. KRIEST locates the IDIM nearest points to the position to be estimated; $IDIM < 16$.)

5. Prompt = DO YOU WISH A LOG TRANSFORM OF Z? ENTER 1 = YES, 0 = NO .
User response = ILOG . (The user must enter a value of ILOG (1 or 0) accordingly. If ILOG equals one then KRIEST analyzes $\ln(Z)$ data. The log transformation can result in a less variable data base and ensures positive predictions for the kriged surface. This latter property is useful for contaminant plume data where negative values are unrealistic.)
6. Prompt = ENTER THE ORDER OF DRIFT . User response = k . (The user must enter the order of drift (k) which corresponds to the degree of polynomial that models the slowly varying trend in the data. Typically, FINDK is run to determine k; k = 0, 1, or 2.)
7. Prompt = ENTER THE GENERALIZED COVARIANCE COEFFICIENTS - (*) FORMAT, CO(DIRAC TERM), A0(H**1 TERM), A1(H**3 TERM), A2(H**5 TERM) . User response = CO, A0, A1, A2 . (The user must enter the coefficients of the generalized covariance (CO, A0, A1, A2) where

$$K(h) = CO\delta + A0h + A1h^3 + A2h^5$$

Typically, the user must run KHREG to determine the coefficients of the polynomial generalized covariance. However, if the user is not too concerned with estimation accuracy and just wants to interpolate observed data, letting the order of drift equal one (k = 1) and the generalized covariance be

$$K(h) = -1.0 h$$

is generally acceptable.)

8. Prompt = ENTER THE NAME OF THE KRIGED SURFACE OUTPUT FILE . User response = KRIGEFIL.DAT . (The user must enter the name of the kriged surface output file; no more than 27 characters. The format of the kriged surface output file is compatible with the CNT plot program.)

9. Prompt = ENTER THE NAME OF THE SIGMA OUTPUT FILE . User response = SIGMAFILE.DAT . (The user must enter the name of the kriging error standard deviation file; no more than 27 characters. The format of the kriging error standard deviation file is also compatible with CNT.)
10. Prompt = ENTER NX,NY,X0,Y0,DX,DY (*)FORMAT . User response = NX, NY, X0, Y0, DX, DY . (The user must enter NX - the number of grid points per row ($NX < 50$), NY - the number of grid points per column ($NH < 50$), X0 - the x-value of the lower left corner of the grid, Y0 - the y-value of the lower left corner of the grid, DX - the grid width for the x-component, and DY - the grid width for the y-component.)

KRIEST echoes the input parameters and data to the line printer and for each point of the user defined rectangular grid, calculates the kriged estimate and the kriging error standard deviation. KRIEST displays the kriged estimates and kriging error standard deviations on a rectangular grid line printer plot and also writes these values to their two respective output files, KRIGEFIL.DAT and SIGMAFILE.DAT. The format of the output files corresponds to the following FORTRAN lines where $D(I,J)$ represents the kriged estimate or kriging standard deviation at the J-th row, I-th column grid point.

```

      DO 10 J=1,NY
10    WRITE(OUTPUTUNIT,11)          (D(I,J),I=1,NX))
11    FORMAT(8F10.4)

```

Thus, the rectangular grid is written row-by-row starting from the bottom and is compatible with the CNT plotting routine.

The kriging error standard deviation may be used as follows: If $\hat{Z}$ is the kriged estimate of a regionalized variable Z and σ_K is the kriging error standard deviation, then under the assumption of normality

Probability [$\hat{Z} - 1.96 \sigma_K < \text{true } Z < \hat{Z} + 1.96 \sigma_K = 95\%$]

i.e., the true value of Z is within $1.96 \sigma_K$ of the kriged estimate $\hat{Z}$ with 95% confidence.

Figure 3.5 shows a sample run of KRIEST using the same set of tritium data (T0667.TMP) as used in the FINDK and KHREG sample cases with the order of drift equal to zero and the generalized covariance given by

$$K(h) = -0.4159 h$$

3.6 PLTWELL

PLTWELL is a display program that will enable the user to plot the observed data as a function of position and then overplot contours generated by the CNT program. The data matrix for the CNT program may be created using the kriging program KRIEST. In this way, irregularly spaced observed data may be contoured.

To execute PLTWELL on the Geosciences Research and Engineering 11/70 computer system:

1. After logging in on the system: User response = RUN PLTWELL .
2. Prompt = ENTER THE TITLE OF THE STUDY . User response = TITLE .
(The user must enter the title of the plot; 27 characters are permitted.)
3. Prompt = ENTER THE NAME OF THE WELL DATA FILE . User response =
WELLDATANAME.DAT . (The user must enter the name of the well data
file; no more than 27 characters. Each of the N records of the input
file contains (X,Y,Z) with a * type format; for example,

*****KRIEST-A KRIGING ROUTINE THAT CALCULATES THE BEST LINEAR UNBIASED ESTIMATES
AND THE ESTIMATED ERROR STANDARD DEVIATION FROM SPATIALLY DISTRIBUTED DATA*****

```

INPUT FILE NAME=10667.FIP
KRIGED SURFACE OUTPUT FILE NAME=10667.KKK
STANDARD DEVIATION OUTPUT FILE NAME=10667.SIG
NUMBER OF OBSERVATIONS= 37
NUMBER OF POINTS TO BE ESTIMATED= 441
NUMBER OF POINTS IN KRIGING GRID= 8
ORDER OF DRIFT= 1
LOG TRANSFORMATION OF Z?(YES=1,NO=0) 1
NUMBER OF POINTS IN X-GRID= 21
NUMBER OF POINTS IN Y-GRID= 21
X ORIGIN= 25.0000
Y ORIGIN= 25.0000
DELTA X= 5.0000
DELTA Y= 5.0000
GENERALIZED COVARIANCE COEFFICIENTS
CO(DIRAC) H H**3 H**5
0.00000 -0.41590

```

DATA VALUES

X	Y	Z
108.2750	38.4000	3.2000
111.0000	50.5000	2.5000
121.2610	58.3080	2.4000
124.0000	63.2000	2.7000
123.2110	62.5650	2.5000
112.7810	62.9000	3.0000
121.0000	66.0000	2.3000
105.6750	70.9000	2.4500
121.0000	79.0000	2.8000
114.5070	82.7110	3.2000
114.5000	82.8690	2.4000
71.0220	83.2100	2.5000
94.0370	83.8150	0.1300
82.6300	84.6020	530.0000
61.0820	86.3330	2.5000
65.8500	87.8250	3.2000
50.7400	87.9530	3.2000
113.7800	89.1480	2.7000
121.0170	89.3040	2.4000
88.6100	89.3890	580.0000
75.6850	92.8090	640.0000
94.4460	94.6650	1400.0000
100.7190	96.0270	380.0000
68.7030	96.6960	21.8000
78.3220	99.5070	2500.0000
99.8250	103.7000	3.2000
108.4300	108.8490	2.1000
75.5740	109.4040	2.1000
86.1240	109.6920	665.0000
97.2750	111.4740	15.0000
66.5370	110.9940	2600.0000
66.9010	114.2740	460.0000
88.3710	115.0930	3.2000
74.3410	115.8600	2.1000
80.8500	119.0000	2.5000
67.0910	119.2060	3.5000
55.1330	118.8400	5.3000

FIGURE 3.5. Sample Run of KRIEST

KRIGED SURFACE AT GRID POINTS

	X=	25,0000	30,0000	35,0000	40,0000	45,0000	50,0000	55,0000	60,0000	65,0000	
	X=	70,0000	75,0000	80,0000	85,0000	90,0000	95,0000	100,0000	105,0000	110,0000	
	X=	115,0000	120,0000	125,0000	*****						
Y= 125,0000	*										
	*	2,0487	2,1687	2,2928	2,3980	3,8966	3,1937	2,6511	2,4036	1,7256	
	*	3,3437	1,1052	1,1192	1,2212	1,8840	1,7990	1,9077	0,9978	0,8217	
	*	0,5995	0,7801	0,6878							
Y= 120,0000	*										
	*	2,1774	2,4255	2,9832	3,3862	3,8478	5,6463	4,5613	7,1241	4,8944	
	*	2,7934	1,5303	1,9667	1,9905	2,1976	2,8999	1,6234	1,2916	1,1118	
	*	1,0196	0,8524	1,3573							
Y= 115,0000	*										
	*	2,1156	2,8857	3,4631	4,3590	6,3073	9,2378	20,0146	57,0002	190,6438	
	*	33,7024	2,4212	4,1259	7,2955	4,8251	6,7096	3,7287	1,9646	1,3960	
	*	1,3604	1,5522	1,5923							
Y= 110,0000	*										
	*	2,2460	2,6576	3,3008	4,4034	8,4172	17,1572	44,9770	199,5755	1168,0253	
	*	167,8839	3,0462	16,0419	307,7250	91,9223	22,0568	6,1630	2,6781	1,8989	
	*	1,5156	1,7457	1,5794							
Y= 105,0000	*										
	*	2,3074	2,7510	3,4556	4,6885	7,1512	13,0362	30,6023	168,4531	329,3208	
	*	133,7376	49,7353	228,0282	712,0599	305,5419	34,8258	3,7436	3,4640	2,7205	
	*	2,3429	1,9092	1,5893							
Y= 100,0000	*										
	*	2,2886	2,6986	3,3296	4,3711	6,2572	10,0425	9,0020	41,2268	46,3933	
	*	74,9007	403,0621	1888,5521	1364,0092	539,7460	231,6195	44,4765	13,5514	5,4065	
	*	3,1938	2,2833	1,7865							
Y= 95,0000	*										
	*	1,8239	2,0727	2,4411	3,0107	3,9210	5,3666	9,3065	10,2902	12,2366	
	*	37,7203	520,9238	1690,0577	1971,1105	1271,6733	1101,3878	326,5242	43,2879	9,7094	
	*	3,8268	2,4269	1,9384							
Y= 90,0000	*										
	*	1,6718	1,8656	2,1326	1,9914	2,5689	3,3553	4,0858	4,0405	4,5278	
	*	16,5305	157,9464	629,1290	1139,0464	363,4327	57,2295	30,3002	15,5982	9,4426	
	*	2,6858	2,4016	1,9042							
Y= 85,0000	*										
	*	1,5231	1,2327	1,4270	1,6923	2,0454	2,4838	2,4118	2,2469	2,2507	
	*	3,7367	29,4195	262,4813	208,8309	9,4699	0,4989	1,3682	3,0623	3,1102	
	*	2,3393	2,5739	2,5380							
Y= 80,0000	*										
	*	1,0107	1,1219	1,2618	1,4312	2,2190	2,0762	1,9446	1,8426	1,9071	
	*	2,9127	7,5913	23,0534	12,5902	1,6146	0,3192	0,5237	1,4580	2,0985	
	*	3,9505	2,8062	2,6563							
Y= 75,0000	*										

FIGURE 3.5. (contd)

	*	0.9439	1.0306	1.1322	2.1056	1.9352	1.8842	1.8571	1.9336	2.5390
	*	4.3448	2.7046	2.8652	3.1135	1.1047	0.8763	0.9286	1.6034	2.7159
	*	2.7781	2.7659	2.6704						
Y= 70.0000	*									
	*	0.8923	2.2857	2.0713	1.8797	1.8693	1.8897	1.9842	2.5208	3.3182
	*	5.5588	1.2006	2.4892	1.3844	0.9946	1.7770	1.4820	2.2555	2.6461
	*	2.6697	2.5665	2.5570						
Y= 65.0000	*									
	*	1.7390	2.0694	1.8594	1.8705	1.9126	2.0132	2.5365	3.0307	4.0996
	*	0.5293	1.3867	1.5433	1.7764	1.1320	3.3724	1.4439	2.6221	2.7930
	*	2.7219	2.3855	2.6154						
Y= 60.0000	*									
	*	2.0813	1.8543	1.8775	1.9290	2.0278	2.5461	2.9055	3.5635	0.8391
	*	0.3624	1.0892	2.0235	2.2064	2.0772	7.0071	2.6987	2.7318	2.8125
	*	2.7141	2.4584	2.5351						
Y= 55.0000	*									
	*	1.8562	1.8862	1.9414	2.0353	2.5488	2.8328	3.2979	0.7139	0.2831
	*	1.5765	1.7003	2.1107	2.1188	12.2167	1.5646	2.7408	2.7021	2.6390
	*	2.5719	2.4857	2.5056						
Y= 50.0000	*									
	*	1.8952	1.9511	2.0392	2.5476	2.7829	3.1396	0.6473	0.6109	1.6902
	*	4.1401	4.3199	2.0497	4.2153	12.0175	2.8349	2.8018	2.7202	2.5551
	*	2.5596	2.5704	2.5901						
Y= 45.0000	*									
	*	1.9590	2.0413	2.5446	2.7456	3.0336	1.0656	1.1785	1.1134	2.5645
	*	4.0164	4.0349	4.1301	4.0414	2.9238	2.9256	2.9146	2.8729	2.7866
	*	2.7223	2.6960	2.6885						
Y= 40.0000	*									
	*	2.0425	2.5410	2.7163	0.8950	0.9440	1.0139	0.9602	1.0569	2.4446
	*	2.3998	3.8966	3.9826	12.7703	2.9997	3.0209	3.0479	3.0875	3.0596
	*	2.9209	2.8387	2.7956						
Y= 35.0000	*									
	*	0.7868	0.8016	0.8240	0.8564	0.9020	0.8626	0.9231	2.2220	2.3431
	*	2.2680	3.8801	3.8727	3.0379	3.0638	3.0961	3.1380	3.1870	3.1667
	*	3.0485	2.9530	2.8908						
Y= 30.0000	*									
	*	0.7535	0.7688	0.7909	0.8217	0.7958	0.8352	2.0711	2.1588	2.0437
	*	2.1637	3.7973	3.0505	3.0824	3.1101	3.1417	3.1743	3.1941	3.1705
	*	3.1013	3.0251	2.9631						
Y= 25.0000	*									
	*	0.7249	0.7402	0.7616	0.7477	0.7738	1.9624	2.0270	2.1016	1.9766
	*	2.1185	2.3419	3.0417	3.1150	3.1402	3.1654	3.1858	3.1913	3.1703
	*	3.1240	3.0672	3.0136						

FIGURE 3.5. (contd)

STANDARD DEVIATION SURFACE AT THE GRID POINTS

		$\lambda = 25.0000$	30.0000	35.0000	40.0000	45.0000	50.0000	55.0000	60.0000	65.0000
	$\lambda = 70.0000$	75.0000	80.0000	85.0000	90.0000	95.0000	100.0000	105.0000	110.0000	
	$\lambda = 115.0000$	120.0000	125.0000	*****						
Y= 125.0000	*									
	*	4.8506	4.4770	4.0750	3.6345	3.1960	2.6220	2.1723	2.0822	2.0192
	*	2.0396	2.1091	2.0794	2.2351	2.4876	2.6813	2.8779	3.0746	3.2999
	*	3.5760	3.8449	4.1450						
Y= 120.0000	*									
	*	4.1054	4.3163	3.8972	3.4280	2.8751	2.1967	1.0627	1.5120	1.2304
	*	1.3344	1.4760	0.9601	1.5561	1.8320	2.1269	2.3429	2.5517	2.7960
	*	3.1020	3.4580	3.8173						
Y= 115.0000	*									
	*	4.5940	4.1985	3.7776	3.3210	2.8087	2.2074	1.6602	1.5974	1.1279
	*	1.1545	0.8659	1.3622	1.2809	1.0656	1.5215	1.7089	1.9629	2.1704
	*	2.6119	3.0858	3.5086						
Y= 110.0000	*									
	*	4.5104	4.1136	3.7020	3.2798	2.8525	2.4512	2.1368	1.9023	1.2453
	*	1.2825	0.7747	1.4226	0.9166	1.3808	1.2750	1.3043	1.4113	1.2527
	*	2.2920	2.7781	3.2242						
Y= 105.0000	*									
	*	4.4493	4.0469	3.6359	3.2314	2.8621	2.5619	2.3300	2.1475	1.8072
	*	1.6133	1.4519	1.5095	1.5767	1.6827	1.5450	0.9531	1.4683	1.6079
	*	2.1165	2.5629	2.9615						
Y= 100.0000	*									
	*	4.4098	3.9932	3.5621	3.1330	2.7470	2.4672	2.2823	2.1106	1.7250
	*	1.4249	1.3459	1.1136	1.7261	1.7454	1.5182	1.2560	1.6085	1.8862
	*	2.0894	2.3292	2.6637						
Y= 95.0000	*									
	*	4.3994	3.9637	3.4948	2.9925	2.4843	2.1138	2.0485	1.9052	1.5072
	*	1.1333	1.1411	1.4317	1.6612	1.4886	0.6959	0.9343	1.5861	1.7915
	*	1.7908	1.8769	2.2391						
Y= 90.0000	*									
	*	4.4283	3.9797	3.4794	2.8984	2.1778	1.2991	1.5846	1.5161	1.2041
	*	1.3978	1.2921	1.4793	1.3788	1.0007	1.4283	1.6542	1.7779	1.5327
	*	1.0089	0.9367	1.7600						
Y= 85.0000	*									
	*	4.5078	4.0639	3.5691	2.9952	2.2951	1.5463	1.6155	1.2351	1.2174
	*	1.0969	1.4534	1.2719	1.1862	1.3422	1.0435	1.7612	1.8809	1.5661
	*	1.0914	1.3741	1.7845						
Y= 80.0000	*									
	*	4.6398	4.2241	3.7737	3.2872	2.7944	2.4022	2.2203	2.0515	1.8860
	*	1.5039	1.7498	1.7991	1.8006	1.7509	1.6008	1.8376	1.8780	1.6685
	*	1.2642	0.9729	1.6864						

FIGURE 3.5. (contd)

Y= 75.0000	*									
	*	4.8201	4.4437	4.0532	3.6672	3.2945	3.0070	2.8142	2.6654	2.5296
	*	2.4218	2.4073	2.4364	2.3307	2.2391	2.0951	1.8771	1.5817	1.6857
Y= 70.0000	*	5.0154	4.7214	4.3758	4.0439	3.7513	3.5162	3.3400	3.2097	3.1067
	*	3.0470	2.9995	2.8535	2.7687	2.6133	2.3910	1.9327	0.9412	1.5157
	*	1.6754	1.5163	1.7517						
Y= 65.0000	*	5.2766	4.9885	4.6899	4.4141	4.1716	3.9718	3.8203	3.7033	3.6221
	*	3.5359	3.3648	3.2760	3.0343	2.8476	2.5907	2.2285	1.7917	1.4013
	*	1.2964	0.9648	1.1918						
Y= 60.0000	*	5.5363	5.2638	5.0067	4.7708	4.5639	4.3958	4.2593	4.1577	4.0442
	*	3.9997	3.7748	3.4663	3.2335	3.0302	2.7662	2.4959	2.0418	1.5471
	*	1.3778	1.1019	1.3475						
Y= 55.0000	*	5.7845	5.5445	5.3191	5.1139	4.9386	4.7880	4.6682	4.5379	4.4748
	*	4.2181	3.9004	3.6162	3.4083	3.2019	2.9343	2.5610	2.0622	1.5849
	*	1.5733	1.5026	1.8932						
Y= 50.0000	*	6.0391	5.8246	5.6249	5.4484	5.2904	5.1577	5.0142	4.9236	4.2464
	*	4.1764	3.9854	3.7935	3.5306	3.2990	3.0467	2.5498	1.9660	0.9304
	*	1.5898	2.0416	2.4111						
Y= 45.0000	*	6.2946	6.1016	5.9274	5.7667	5.6263	4.9765	4.7998	4.6418	4.5045
	*	4.3829	4.1742	3.9395	3.6572	3.5495	3.0568	2.5080	1.9260	1.5980
	*	1.9133	2.3675	2.7711						
Y= 40.0000	*	6.5490	6.3785	6.2180	5.5867	5.3958	5.2114	5.0425	4.8651	4.7138
	*	4.5763	4.3525	4.1027	3.8839	3.6263	3.1069	2.4822	1.6461	1.2829
	*	2.0798	2.6229	3.0586						
Y= 35.0000	*	6.3996	6.1965	5.9987	5.8063	5.6189	5.4433	5.2628	5.1073	4.9186
	*	4.7725	4.5635	4.2847	4.2169	3.7721	3.2697	2.6781	1.9753	1.7690
	*	2.3875	2.9152	3.3455						
Y= 30.0000	*	6.5922	6.3971	6.2061	6.0189	5.8398	5.6590	5.5013	5.3172	5.1994
	*	4.9758	4.7680	4.4903	4.4044	3.9979	3.5628	3.1135	2.7367	2.6583
	*	2.9249	3.3000	3.6691						
Y= 25.0000	*	6.7822	6.5941	6.4091	6.2290	6.0499	5.8914	5.7121	5.5283	5.4093
	*	5.1842	4.9899	4.9842	4.6400	4.2861	3.9331	3.6109	3.3851	3.3383
	*	3.4806	3.7340	4.0269						

FIGURE 3.5. (contd)

2.50, 3.50, 457.63

or

-----2.50-----3.50-----457.63

where - indicates a
space.)

4. Prompt = ENTER THE NUMBER OF WELLS . User response = N . (The user must enter the number (N) of ordered triples in the observed data base disk file; $N \leq 100$. An ordered triple is (X,Y,Z) where (X,Y) is the spatial position and Z is the observed value.)
5. Prompt = ENTER PLOTFILE NAME FOR WELL DATA . User response = WELLDATA.PLT . (The user must enter the plotfile name; 27 characters are allowed.)
6. Prompt = ENTER NX,NY,DX,DY,X0,Y0(*)FORMAT . User response = NX, NY, DX, DY, X0, Y0 . (The user must enter NX - the number of grid points per row ($NX \leq 50$), NY - the number of grid points per column ($NY \leq 50$), DX - the grid width for the x-component, DY - the grid width for the y-component, X0 - the x-value of the lower left corner of the grid, and Y0 - the y-value of the lower left corner of the grid.)
7. Prompt = ENTER THE NUMBER OF DATA UNITS PER INCH FOR PLOTS . User response = INCH . (The user must supply this value (INCH) according to the range of x-y values in the data base.)
8. Prompt = ENTER THE NUMBER OF DECIMAL POINTS ON LABELS . User response = IDPTS . (The user must supply this value (IDPTS); negative integer value indicates decimal points are not to be written on labels.)
9. Prompt = ENTER CNT.CTL NAME . User response = CONTROLFILE.CTL . (PLTWELL creates a control file for CNT that enables the user to overplot contours on the plot produced by WELLDATA.PLT (of step 5).

10. Prompt = ENTER CONLOW,CONHGH,CONDIF . User response = CONLOW, CONHGH, CONDIF . (The user must enter the lowest contour value (CONLOW), highest contour value (CONHGH) and difference between contour values (CONDIF) as input to CNT.)
11. Prompt = ENTER MATRIX INPUT FILE NAME . User response = MATRIXINPUT.DAT . (The user must enter the name of the data matrix file to be used for contouring in CNT; 27 characters are permitted. The matrix input file format corresponds to the following FORTRAN lines where D(I,J) represents the parameter value at the J-th order, I-th column grid point.

```
      DO 10 J=1,NY
10    WRITE (MATRIXINPUT,11) (D(I,J),I=1,NX)
11    FORMAT(8F10.4)
```

This format is exactly the same as the KRIEST output files.)

PLTWELL echoes the input parameter values to the line printer and creates a device-independent graphics plot file (DIG file) named WELLDATA.PLT (user selected name) and a CNT control file named CONTROLFILE.CTL (user selected name) to permit contour overplotting. CNT may then be executed using CONTROLFILE.CTL to produce a second plot file that will be named CONTOUR.PLT for discussion purposes. WELLDATA.PLT utilizes pen number two; it is advisable to use black or dark blue ink. CONTOUR.PLT utilizes pen number one; it is advisable to use red ink for the overplotted contours. NOTE: To overplot successfully on the CALCOMP plotter the origin on the plotter must be moved to coincide with pen number one after plotting WELLDATA.PLT with pen number two. Otherwise, the overplot is offset approximately one-half inch throughout the plot.

Figures 3.6 and 3.7, respectively, show a sample run and plot of PLTWELL. The plot includes CNT overplots using the same set of tritium data (T0667.TMP) from the FINDK, KHREG, and KRIEST sample runs and the kriged estimate output file from KRIEST (T0667.KKK).

*****PLTWELL-- A PROGRAM TO PLOT WELL LOCATIONS ON A CMT CONTOUR PLOT

WELL DATA FILE=T0667.DAT
 NUMBER OF WELLS= 37
 NAME OF WELL DATA PLOT FILE=T0667.PLT
 NUMBER OF ROWS ON X AXIS= 21
 NUMBER OF ROWS ON Y AXIS= 21
 X GRID SIZE= 10.0000
 Y GRID SIZE= 10.0000
 X ORIGIN= 25.0000
 Y ORIGIN= 25.0000
 THE NUMBER OF DATA POINTS PER INCH= 15.0000
 THE NUMBER OF DECIMAL POINTS ON LABELS= 0

DATA VALUES X-Y-Z		
108.2750	38.40000	3.200000
111.0000	50.50000	2.500000
121.2610	58.30800	2.400000
124.0000	63.20000	2.700000
123.2110	62.56500	2.500000
112.7810	62.90000	3.000000
121.0000	66.00000	2.300000
105.6750	70.90000	2.450000
121.0000	79.00000	2.800000
114.5070	82.71100	3.200000
114.5000	82.86900	2.400000
71.02200	83.21000	2.500000
94.03700	83.81500	0.1300000
82.63000	84.60200	530.0000
61.68200	86.33300	2.500000
65.85000	87.82500	3.200000
50.74000	87.95300	3.200000
113.7800	89.14800	2.700000
121.0170	89.30400	2.400000
88.61000	89.38900	580.0000
75.68500	92.80900	640.0000
94.44600	94.66500	1400.000
100.7190	96.02700	380.0000
68.70300	96.69600	21.80000
78.32200	99.50700	2500.000
99.82500	103.7000	3.200000
108.4300	108.8490	2.100000
75.57400	109.4040	2.100000
86.12400	109.6920	665.0000
97.27500	111.4740	15.00000
66.93700	110.9940	2600.000
66.90100	114.2740	460.0000
88.37100	115.0930	3.200000
74.34100	115.8600	2.100000
80.65000	119.0000	2.500000
67.09100	119.2000	3.500000
55.73300	118.8400	5.300000

CMT.CTL FILE NAME=T0667.CTL
 MATRIX INPUT FILE FOR CMT=T0667.KKK

FIGURE 3.6. Sample Run of PLTWELL

TRITIUM PLUME 4-6/67

FIGURE 3.7. Plot from PLTWELL with Tritium Contours for 1967

4.0 BIBLIOGRAPHY

- Chiles, J. P. 1977. Geostatistique des Phenomenes Non Stationnaires (in French), These de Docteur-Ingenieur, University de Nancy, Nancy, France.
- Delfiner, P., and J. P. Chiles. 1978. "Conditional Simulations: A New Monte Carlo Approach to Probabilistic Evaluation of Hydrocarbon in Place." Ecole des Mines de Paris, Fontainebleau, France.
- Gambolati, G., and G. Volpi. 1979. "Groundwater Contour Mapping in Venice by Stochastic Interpolators. 1. Theory." Water Resources Research 15(2):281-290.
- Gelb, A. 1974. Applied Optimal Estimation. M.I.T. Press, Cambridge, Massachusetts.
- Matheron, G. 1965. Les Variables Regionalisees et leur Estimation. Masson and Cie, France.
- Matheron, G. 1967. Elements pour une Theorie des Milieux Poreus. Masson and Cie, France.
- Matheron, G. 1969. "Lew Krigeage Universel." Les Cashiers du CMM; fasc. 1., Fontainebleau, France.
- Mejia, J. M. 1974. "On the Synthesis of Random Field Sampling from the Spectrum: An Application to the Generation of Hydrologic Spatial Processes." Water Resources Research 10(4):705-711.
- Volpi, G., and G. Gambolati. 1978. "On the Use of a Main Trend for the Kriging Technique in Hydrology." Advances in Water Resources 1(6):345-349.

DISTRIBUTION

No. of
Copies

OFFSITE

27 DOE Technical Information Center

ONSITE

3 DOE Richland Operations Office

R. E. Austin
D. R. Elle
H. E. Ransom

3 Rockwell Hanford Operations

R. C. Arnett
R. G. Baca
V. Johnson

32 Pacific Northwest Laboratory

P. E. Bramson
C. S. Cline
J. L. Devary (5)
D. W. Damschen (2)
P. A. Eddy
E. L. Hilty
T. L. Liikala
V. L. McGhan
P. J. Mitchell (2)
C. A. Newbill
L. S. Prater (2)
K. R. Price
J. R. Raymond
A. E. Reisenauer
W. A. Rice (2)
J. T. Rieger
P. D. Thorne
Publishing Coordination (2)
Technical Information (5)

.

.

.

.

.

.

.

.